

FP7 Grant Agreement

266632

Deliverable No and Title D3.6 – Web 2.0 guidelines

Dissemination Level PU (public)

Work Package WP3

Version 1

Release Date 28 February 2014

Author(s) Judit Bar-Ilan, Lorna Wildgaard

**Project Website** 

http://research-acumen.eu/

European Commission 7th Framework Programme SP4 - Capacities Science in Society 2010

**Grant Agreement: 266632** 





# D3.6 Web 2.0 guidelines + guidelines for searching citation databases

The first part of this deliverable provides best practice guidelines for researchers engaging with Web 2.0 and web media.

## Part 1

We recommend the following Web 2.0 practices:

Academics should make sure that they are visible on the Web and their work is accessible. A stable, easy to find Web presence is essential for researchers. This can be achieved by the following:

- Web CV in English and full list of publications on the personal website preferably maintained on the servers of the institution where the academic works. The list of publications should be updated periodically.
- 2. Uploading OA publications or OA versions of publications ("green open access", see <a href="http://www.sherpa.ac.uk/documents/sherpaplusdocs/Nottingham-colour-guide.pdf">http://www.sherpa.ac.uk/documents/sherpaplusdocs/Nottingham-colour-guide.pdf</a>) of the academic either to the personal website, or to the institutional repository or to a subject specific repository if such a repository exists (e.g. arXiv for astrophysicists or philpapers.org for philosophers). If there is a well-known subject specific repository, then submitting to this repository should be considered a "must".
- 3. Maintaining a presence on professional social media sites: academia.edu and ResearchGate, and on the more general professional social media site, LinkedIn (it should be noted that these are our current best practices, social media sites might disappear and others can appear in their place). These social media sites allow other members of the network to follow the researcher and to get update on his/her activities.
- 4. One of the major problems for evaluation is disambiguation of author names. WoS offers the ResearcherID and Scopus the Scopus Author ID to solve this problem, but perhaps a better solution is the newly introduced ORCID which is able to combine publication lists from different sources and "supports automated linkages between you and your professional activities ensuring that your work is recognized" (<a href="http://orcid.org">http://orcid.org</a>). Creating and maintaining an ORCID identifier is quite simple.
- 5. Another option to enhance visibility is to set up a Google Scholar Citation page, which allows cleaning the citation and publication data provided by Google Scholar (see <a href="http://www.google.com/intl/en/scholar/citations.html">http://www.google.com/intl/en/scholar/citations.html</a>).
- 6. Presentations can be uploaded to slideshare (<a href="http://www.slideshare.net/">http://figshare.com/</a>) and videos of presentations to youtube or vimeo (<a href="https://vimeo.com/">https://vimeo.com/</a>), although these are often available on the event site as well. The specific sites mentioned are not that important, as long as the researcher's output are available and visible online. Visibility can be increased by linking to all the available materials from the researcher's homepage and from other profiles the researcher maintains on social media sites.

- 7. Academics may also consider maintaining science blogs, in which case it is recommended to register with aggregators, (e.g. researchblogging.org or scienceseeker.org) to publicize blog posts. However, the academic should keep in mind that if the blog is not regularly updated it is better not to have a blog at all. If the researcher knows in advance he/she does not have time to maintain his/her own blog, an option is to guest-blogging in other science blogs or writing in a group blog.
- 8. Twitter is also an option for disseminating information, however currently even though all tweets are deposited at the Library of Congress, they are practically inaccessible, and thus currently tweets should be considered as ephemera, however they can cause momentary hype and thus increase the researcher's visibility.

## Part 2

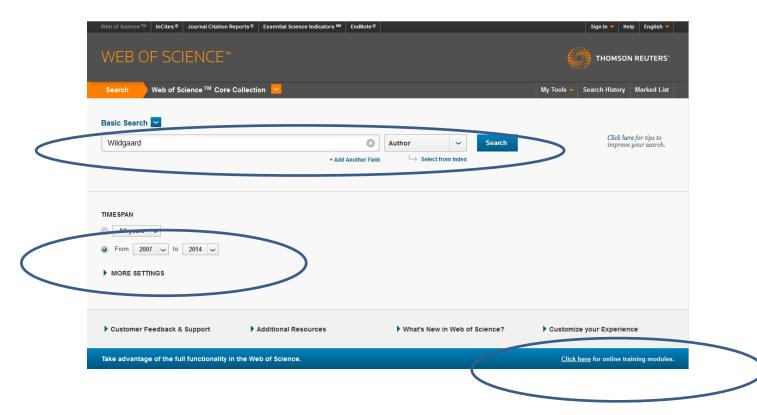
In addition to the Web 2.0 guidelines, we provide detailed instructions on searching for publications and citation counts on Thomson Reuters' Web of Science, on Elsevier's Scopus and on Google Scholar; for retrieving readership counts on Mendeley, for retrieving the number of downloads of publications from specific social media sites. We decided to include these instructions in this deliverable, since they are essential for filling in the ACUMEN Portfolio. Always report the exact date and source of data collection.

Note that these instructions are accurate as of February 2014. There are frequent changes to the search interfaces.

# Web of Science (WOS)

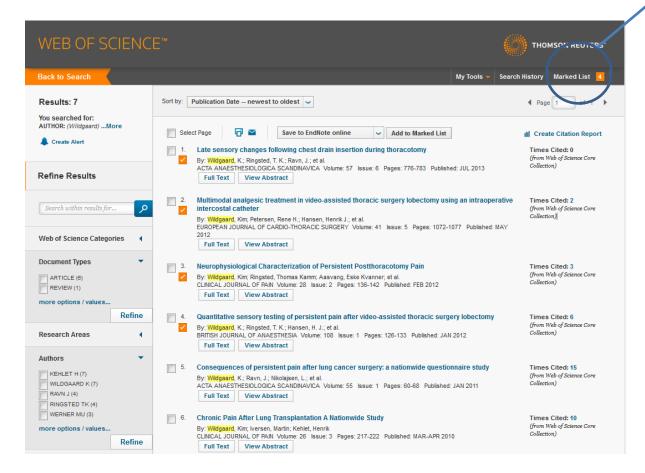
## Retrieve full list of publications indexed by this database

Open Web of Science (a citation database that is part of Web of Knowledge). You need subscription to access WOS)



Enter your surname and initials in the search box. Limit the field to "author". Limit the search, under Timespan, from the publication year of your first publication. Another option is to search each title of your publications one by one. Limit in this case the field to "title".

Press search. (Please explore the training modules for online tutorials)



First limit to author name: In the column refine results click on Author, and more options. Click the surname and initial option(s) that are relevant and click refine to just include these variants.

Repeat for Web of Science categories if necessary. If there are just a few categories click on those you wish to exclude and then click on "exclude". If there are many options, select the relevant categories and "refine". Think broadly when using the categories and narrow the search slowly, continuously checking the results list. Philosophy can for example also be included in the mathematics, social studies, or management category. This is especially important if you have a common name.

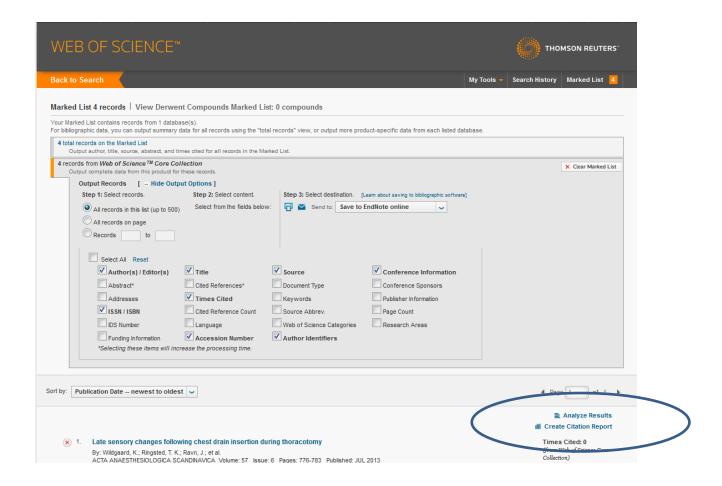
When you narrowed the list with the above-mentioned refinements, click the boxes beside the items to add your articles to the marked list. You find the marked list at the top of the search.

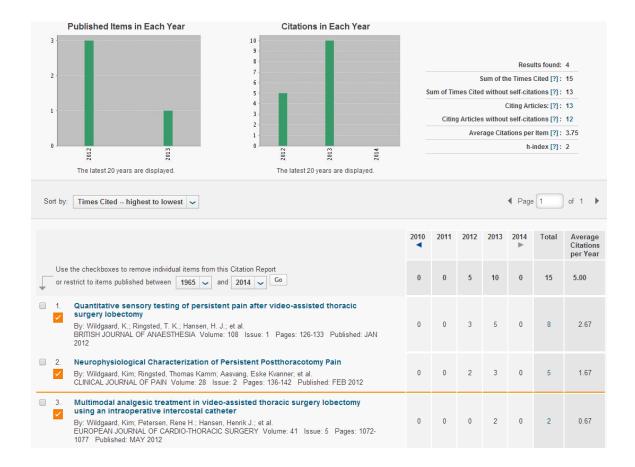
Click on the number in the orange box to enter your marked list. In the marked list you can remove items that were not authored by you or return to the search to find and add more publications. When you are satisfied with your list of publications, you can "create a citation report".

## Create a citation report

Enter your marked list, and click on "create citation report". You can now read the total number of your publications indexed by WOS, the total number of times cited, the average number of

citation per paper and your h-index. You can also see a histogram of your publications over time as well as citations over time.

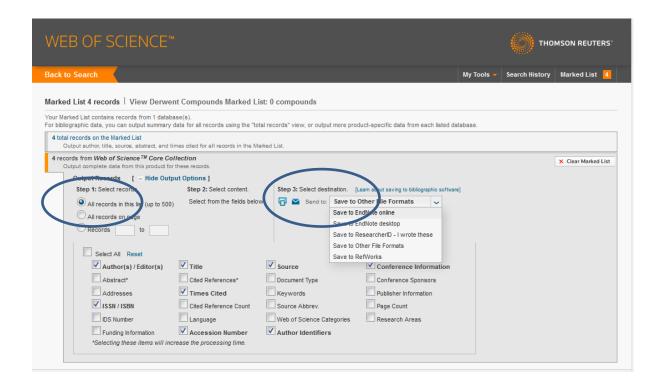




## How to export your publication and citation data

(This is useful for supporting the data in the portfolio, as citation and publication data changed over time)

- Step 1: in the marked list check "All records in this list"
- Step 2: select the information you wish to export.
- Step 3: Select the destination. You can choose reference managers Endnote or Refworks, save to your Researcher ID or choose "Other Formats". Choosing "other formats" allows you to save in a Tab de-limited Win or Mac file so you can export to an excel file on your computer.



# **Google Scholar**

## **Google Scholar Citation Profile**

If you established a Google Scholar Citation profile, check for the accuracy of the list, and then you will have the information you need for the ACUMEN Portfolio. Please consult the help pages to understand how to include/exclude certain publications (see <a href="http://scholar.google.com/intl/en/scholar/citations.html#setup">http://scholar.google.com/intl/en/scholar/citations.html#setup</a>). There is an option to export your publications, by choosing Export from the Actions menu (available only to the owner of the profile); however this will not export the number of citations each publication received. You can add this information manually, but unfortunately the saved records do not appear in the same order as the displayed record.

Another option to save the information on the GS output as evidence of the information provided in the portfolio, is to save the html file using the "save" or "save as" options of your web browser. GS displays 20 results per page, you have to save each 20 separately (giving each saved file a different name). You can browse forward by clicking next.

#### **Publish or Perish**

If you do not have a GS Citation Profile or it not complete, use Publish or Perish.

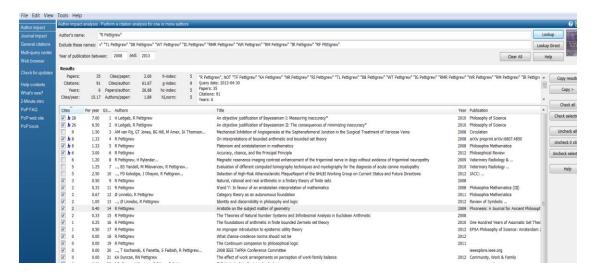
#### Download and install Publish and Perish: http://www.harzing.com/pop.htm

Search using the **Author Impact** function. The Author impact analysis page allows you to perform a quick analysis of the impact of your publications. The Author impact analysis page contains the following panes:

- -Author query pane
- -Results pane

#### How to perform an Author impact analysis

- 1. Enter the your name in the Author's name field;
- 2. Click Lookup or press the Enter key.
- 3. The program will now contact Google Scholar to obtain the citations, process the list, and calculate the Citation metrics, which are then displayed in the Results pane. The full list of results is also available for inspection or modifications and can be exported in a variety of formats.



From your publication list see how your name is written in the author byline. Use these forms to search the databases. Write them with "quotes" with OR in between each name:

"Piotr A Dybczynski" OR "PA Dybczynski" OR "Dybczynski, P"

Search tips (based on <a href="http://www.harzing.com/pophelp/author-impact.htm">http://www.harzing.com/pophelp/author-impact.htm</a>)

1. Always use "quotes" around the author's name, e.g. "L Wildgaard". Opt for initials instead of full given name, as not all journals publish author names in full.

- 2. Another approach is to search the title of each of your publications one-by-one. Use "General Citations" and search the title in the field "all of the words".
- 3. PoP is not case sensitive.
- 4. The order of search terms does not matter. "L Wildgaard" will give the same result as "Wildgaard L".
- 5. If you have consistently published with only one initial, you can exclude namesakes using 2nd and 3rd initials by using wildcards in the "exclude these names" field, e.g. when searching for "L Wildgaard", you can exclude "L\* Wildgaard" "L\*\* Wildgaard".
- 6. You cannot use "\*L Wildgaard" to exclude "KL Wildgaard". You need to manually exclude these authors by listing them in the "exclude these names" field. To exclude certain author names, enter them in the Exclude these names field. You can enter more than one exclusion in Exclude these names.
- 7. If an author has published under two different names (e.g. maiden name and married name) use OR between the names under which you published.
- 8. If your name contains accented letters, then it might help if you include several variations of the name, both with and without accents, and also with the accented letters missing. For example, to search for someone with the surname Veríssimo (note the accent on the first 'i'), use the following names in the Author field: "Veríssimo" OR "Verissimo" OR "Verssimo"
- 9. If the list of results is fairly limited, you can manually include or exclude citations from the analysis by checking or clearing the boxes in the Results list.

#### **Limiting year**

Before limiting the year range, always check whether you have publications without a year listing. If you know that you only published after (or before) a certain year, you can enter the start or end years in the Year of publication between ... and ... fields.

#### (De)Selecting and merging results

You can deselect publications not published by the you. Simply remove the tick mark in the first column by clicking on it.

You can (de)select more than one publication at once by first selecting the relevant publications and then clicking the "(un)check selection" button.

If the results contain duplicate entries, you can merge them by dragging and dropping the duplicate entries onto the master record.

Selecting relevant publications for unchecking or merging can be made easier by first sorting the results by Cites, Authors, Title, Year, Publication, or Publisher. Sorting is done simply by clicking on the corresponding column heading. Click twice to reverse the sort order.

#### Here are some shortcuts:

- 1. The Check all button places check marks in all boxes;
- 2. The Uncheck all button clears all boxes;
- 3. When you use the keyboard to move up and down in the Results list, pressing the space bar toggles the check mark on and off on the selected line.
- 4. You can also select a consecutive range of items in the list (left-click on the first item, then hold either Shift key and left-click on the last item) and use the Check selection/Uncheck selection buttons to check/uncheck all selected items and recalculate the citation statistics.

Selecting relevant publications might be easier by sorting the results by Cites, Authors, Title, Year, Publication, or Publisher. Sorting is done simply by clicking on the corresponding column heading.

Searching and making the results accurate is time-consuming as in February 2013 Google Scholar reduced the maximum number of results per page from 100 to 20. This has following effect on data extraction (from <a href="http://www.harzing.com/pophelp/faq.htm#Q1010">http://www.harzing.com/pophelp/faq.htm#Q1010</a>):

- More page requests mean that POP hits the maximum number of requests that Google Scholar allows per hour sooner.
- If the number of page requests exceeds the maximum that Google Scholar allows, the IP address will be temporarily blocked by Google Scholar. This block can last for up to 24 hours.
- To avoid hitting the maximum allowable request limit, POP uses an adaptive request rate limiter. This limits the number of requests that are sent to Google Scholar within a given period, both short-term (during the last 60 seconds) and medium term (during the last hour).
- To achieve the required reduction in requests, Publish or Perish delays subsequent requests for a variable amount of time (up to 1 minute). The higher the recent request rate, the longer the delays.

This means the amount of data collection per session is limited and the speed of data extraction can get slower and slower. There is a possibility of being blocked by Google Scholar for up to 24 hours. If you perform queries that yield many results (several hundred or more) and issue a number of queries in short succession, the request rate limiter will insert progressively longer delays to keep the overall request rate within acceptable limits. To avoid this, spread the queries over the day.

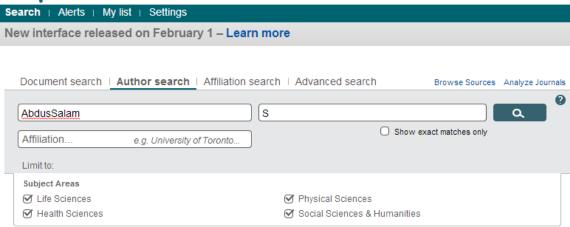
#### **Saving the results:**

POP allows you to save the selected publications together with the citation counts in a CSV format. This option can be selected from the File menu.

# **Scopus**

Open Scopus (<u>www.scopus.com</u>) (requires subscription). Click on author search. Enter your last name and first initial (same as before, if you enter your full given name, you may miss some of the publications authored by you). You can also limit to institution, but this will only work if you published under the same affiliation during your whole career.

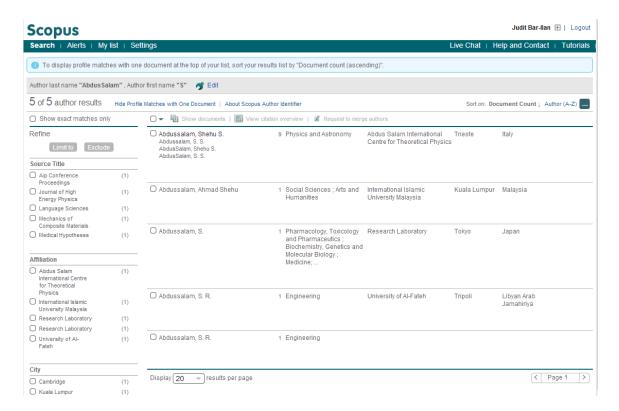
# Scopus



See results for the search AbduSalam S



It says that there are 5 results, but by default Scopus only shows results with two documents or more. Click on "Show Profile Matches with One Document" to see all the versions. If there are many results, it is useful to change the number of displayed results per page to a larger number (the maximum is 200) in order to be able to view all the results.

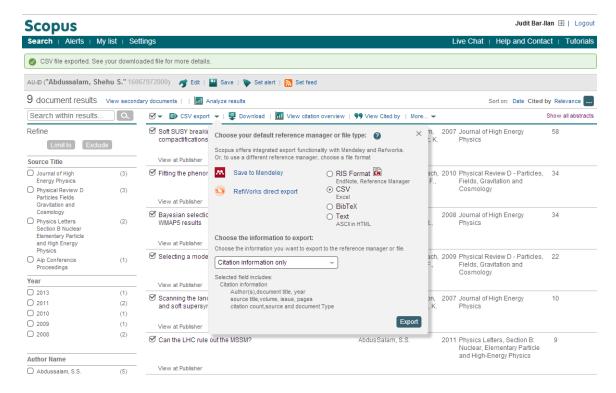


Select all relevant records (click on the box beside the name) – if there is more than one relevant record, click all the relevant boxes)

Click on show documents (on top). Sort the results by "Cited by" (this option is on the top right). Click "All" on the top left. Here too you might want to change the number of displayed records per page to 200, check to see if all the docs are authored by you. Unselect those that were not. You may use the filtering options on the left side of the screen to exclude publications.



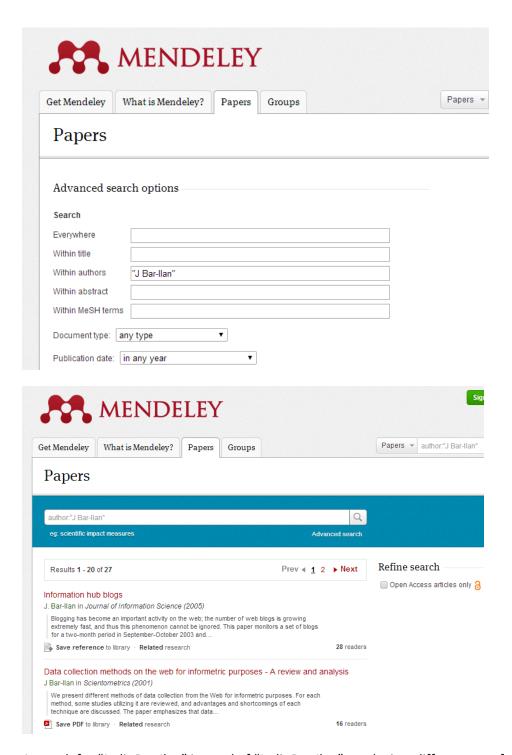
Then click on "Export". Choose Export format as: CSV + citation information only. Here too you have to search for all the name variations (initials, maiden name etc.) under which you published.



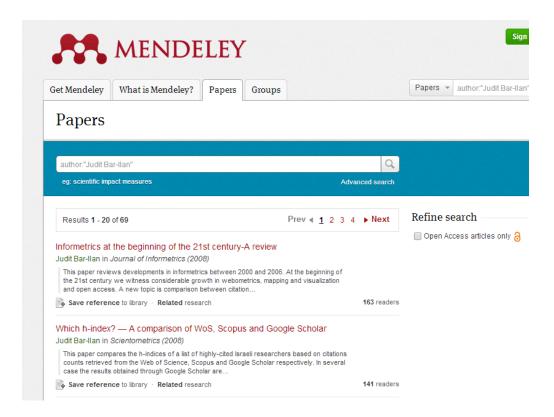
If you click on "View citation overview", your h-index according to Scopus will be displayed. You will also see the total number of publications and the total number of citations.

# **Mendeley**

Here you want to retrieve the readership counts for your publications with the most "readers" on Mendeley. We suggest starting from the advanced search page (<a href="http://www.mendeley.com/research-papers/search/">http://www.mendeley.com/research-papers/search/</a>). There are two options, you can either search for your name in the authors field (you will have to search for each name variation separately – initials only plus last name, first name-middle initials-last name, first name-middle name(s)-last name, and all the last names under which you published). Currently there is no author disambiguation at all, and the items do not appear by decreasing number of readers, so you will have to scan through all results, to locate the items with the largest number of readers. In the advanced search form you can change the number of results per page to 100.



A search for "Judit Bar-Ilan" instead of "Judit Bar-Ilan" results in a different set of publications.



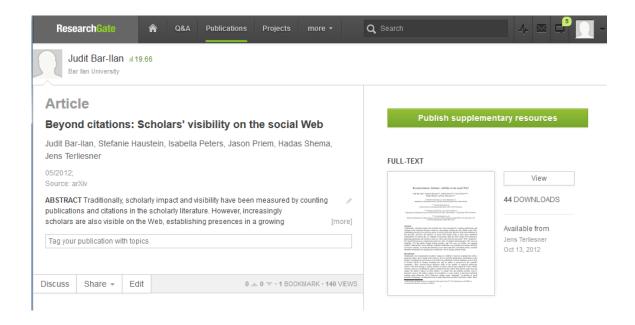
The other option is to search for the title of publications (using quotation marks around the title) you "suspect" are highly read. Mendeley does not handle very well punctuation marks and special characters, we suggest avoiding these and searching for partial titles without special characters instead.

## **Download information**

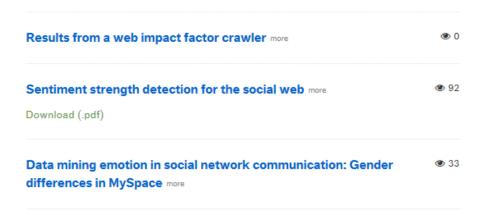
Some journals (e.g. the PLoS journals) provide download counts. Others do not provide exact download counts, but list the top downloaded articles. Some subject-specific repositories, e.g. the Social Science Research Network (SSRN) (<a href="http://www.ssrn.com/en/">http://www.ssrn.com/en/</a>) or philpapers.org also provide download counts of items downloaded from these sites.

You can also get the number of downloads of your papers from academia.edu or ResearchGate.

On ResearchGate when clicking on each of your publications listed on this site, information on the number of downloads is provided.



Academia.edu provides information on the number of document views from this site for each publication listed under your profile.



Note that download information is always only from the specific site.

The number of followers on the academic social media sites and on Twitter can also be easily located.

# **Summary**

In part 1 of this deliverable we provided guideline son how to ensure and increase visibility on the Web. In the second part we gave detailed instructions on how to use different data sources for filling in the ACUMEN Portfolio.